

Berry-Esseen's central limit theorem for non-causal linear processes in Hilbert space

Mohamed EL MACHKOURI

July 7, 2010

Abstract

Let H be a real separable Hilbert space and $(a_k)_{k \in \mathbb{Z}}$ a sequence of bounded linear operators from H to H . We consider the linear process X defined for any k in \mathbb{Z} by $X_k = \sum_{j \in \mathbb{Z}} a_j(\varepsilon_{k-j})$ where $(\varepsilon_k)_{k \in \mathbb{Z}}$ is a sequence of i.i.d. centered H -valued random variables. We investigate the rate of convergence in the CLT for X and in particular we obtain the usual Berry-Esseen's bound provided that $\sum_{j \in \mathbb{Z}} \|a_j\|_{\mathcal{L}(H)} < +\infty$ and ε_0 belongs to L_H^∞ .

Short title: Berry-Esseen's CLT for Hilbertian linear processes.

Key words: Central limit theorem, Berry-Esseen bound, linear process, Hilbert space.

1 Introduction and notations

Let $(H, \|\cdot\|_H)$ be a separable real Hilbert space and $(\mathcal{L}, \|\cdot\|_{\mathcal{L}(H)})$ be the class of bounded linear operators from H to H with its usual uniform norm. Consider a sequence $(\varepsilon_k)_{k \in \mathbb{Z}}$ of i.i.d. centered random variables, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with values in H . If $(a_k)_{k \in \mathbb{Z}}$ is a sequence in \mathcal{L} , we define the (non-causal) linear process $X = (X_k)_{k \in \mathbb{Z}}$ in H by

$$X_k = \sum_{j \in \mathbb{Z}} a_j(\varepsilon_{k-j}), \quad k \in \mathbb{Z}. \quad (1)$$

If $\sum_{j \in \mathbb{Z}} \|a_j\|_{\mathcal{L}(H)} < \infty$ and $E\|\varepsilon_0\|_H < +\infty$ then the series in (1) converges almost surely and in $L_H^1(\Omega, \mathcal{A}, \mathbb{P})$ (see Bosq [2]). The condition $\sum_{j \in \mathbb{Z}} \|a_j\|_{\mathcal{L}(H)} < \infty$ is known to be sharp

for the \sqrt{n} -normalized partial sums of X to satisfies a CLT provided that $(\varepsilon_k)_{k \in \mathbb{Z}}$ are i.i.d. centered having finite second moments (see Merlevede et al. [6]). In this work, we investigate the rate of convergence in the CLT for X under the condition

$$\sum_{j \in \mathbb{Z}} |j|^\tau \|a_j\|_{\mathcal{L}(H)} < \infty \quad (2)$$

with $\tau = 1$ when $(\varepsilon_k)_{k \in \mathbb{Z}}$ are assumed to be i.i.d. centered and such that ε_0 belongs to L_H^∞ and $\tau = 1/2$ when $(\varepsilon_k)_{k \in \mathbb{Z}}$ are i.i.d. centered and such that ε_0 belongs to some Orlicz space $L_{H,\psi}$ (see section 2). This problem was previously studied (with $\tau = 1$ in Condition (2)) by Bosq [3] for (causal) Hilbert linear processes but a mistake in his proof was pointed out by V. Paulauskas [7]. However, in the particular case of Hilbertian autoregressive processes of order 1, Bosq [4] obtained the usual Berry-Esseen inequality provided that $(\varepsilon_k)_{k \in \mathbb{Z}}$ are i.i.d. centered with ε_0 in L_H^∞ .

2 Main result

In the sequel, C_{ε_0} is the autocovariance operator of ε_0 , $A := \sum_{j \in \mathbb{Z}} a_j$ and A^* is the adjoint of A . For any sequence $Z = (Z_k)_{k \in \mathbb{Z}}$ of random variables with values in H we denote

$$\Delta_n(Z) = \sup_{t \in \mathbb{R}} \left| \mathbb{P} \left(\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^n Z_k \right\|_H \leq t \right) - \mathbb{P} (\|N\|_H \leq t) \right|$$

where $N \sim \mathcal{N}(0, AC_{\varepsilon_0}A^*)$.

For any $j \in \mathbb{Z}$, denote $c_{j,n} = \sum_{i=1}^n b_{i-j}$ where $b_i = a_i$ for any $i \neq 0$ and $b_0 = a_0 - A$.

Lemma 1 *For any positive integer n ,*

$$\sum_{k=1}^n X_k = A \left(\sum_{k=1}^n \varepsilon_k \right) + Q_n + R_n$$

where $Q_n = \sum_{k=1}^n \sum_{|j| > n} a_{k-j}(\varepsilon_j)$ and $R_n = \sum_{|j| \leq n} c_{j,n}(\varepsilon_j)$.

Recall that a Young function ψ is a real convex nondecreasing function defined on \mathbb{R}^+ which satisfies $\lim_{t \rightarrow +\infty} \psi(t) = +\infty$ and $\psi(0) = 0$. We define the Orlicz space $L_{H,\psi}$ as the space of H -valued random variables Z defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such

that $E[\psi(\|Z\|_H/c)] < +\infty$ for some $c > 0$. The Orlicz space $L_{H,\psi}$ equipped with the so-called Luxemburg norm $\|\cdot\|_\psi$ defined for any H -valued random variable Z by

$$\|Z\|_\psi = \inf \{ c > 0 ; E[\psi(\|Z\|_H/c)] \leq 1 \}$$

is a Banach space. In the sequel, $c(N)$ denotes a bound of the density of $\mathcal{N}(0, AC_{\varepsilon_0}A^*)$ (see Davydov et al. [5]). Our main result is the following.

Theorem 1 *Let $(\varepsilon_k)_{k \in \mathbb{Z}}$ be a sequence of i.i.d. centered H -valued random variables and let X be the Hilbertian linear process defined by (1).*

i) If ε_0 belongs to L_H^∞ and $\sum_{j \in \mathbb{Z}} |j| \|a_j\|_{\mathcal{L}(H)} < \infty$ then

$$\Delta_n(X) \leq \frac{c_1}{\sqrt{n}} \quad (3)$$

where $c_1 = c_2 + 14c(N)\|\varepsilon_0\|_\infty \sum_{j \in \mathbb{Z}} |j| \|a_j\|_{\mathcal{L}(H)}$ and c_2 is a positive constant which depend only on the distribution of ε_0 .

ii) If ψ is a Young function then

$$\Delta_n(X) \leq \Delta_n(A(\varepsilon)) + \varphi \left(\frac{c(N)\|Q_n + R_n\|_\psi}{\sqrt{n}} \right) \quad (4)$$

where $\varphi(x) = xh^{-1}(1/x)$ and $h(x) = x\psi(x)$ for any real $x > 0$.

The inequality (4) ensures a rate of convergence to zero for $\Delta_n(X)$ as n goes to infinity provided that $\Delta_n(A(\varepsilon_0))$ goes to zero as n goes to infinity and a bound for $\|Q_n + R_n\|_\psi$ exists. As just an illustration, we have the following corollary.

Corollary 1 *Assume that $(\varepsilon_k)_{k \in \mathbb{Z}}$ are i.i.d. centered H -valued random variables and that the condition (2) holds with $\tau = 1/2$.*

i) If ε_0 belongs to L_{H,ψ_1} then $\Delta_n(X) = O\left(\frac{\log n}{\sqrt{n}}\right)$ where ψ_1 is the Young function defined by $\psi_1(x) = \exp(x) - 1$.

ii) If ε_0 belongs to L_H^r for $r \geq 3$ then $\Delta_n(X) = O\left(n^{-\frac{r}{2(r+1)}}\right)$.

3 Proofs

Proof of Lemma 1. For any positive integer n , we have

$$\begin{aligned}
R_n &= \sum_{j=-n}^n c_{j,n}(\varepsilon_j) = \sum_{k=1}^n \sum_{j=-n}^n b_{k-j}(\varepsilon_j) \\
&= \sum_{k=1}^n \sum_{j \in [-n,n] \setminus \{k\}} a_{k-j}(\varepsilon_j) + (a_0 - A) \left(\sum_{k=1}^n \varepsilon_k \right) \\
&= \sum_{k=1}^n \sum_{j=-n}^n a_{k-j}(\varepsilon_j) - A \left(\sum_{k=1}^n \varepsilon_k \right) \\
&= -Q_n + \sum_{k=1}^n X_k - A \left(\sum_{k=1}^n \varepsilon_k \right).
\end{aligned}$$

The proof of Lemma 1 is complete.

Proof of Theorem 1. Let $\lambda > 0$ and $t > 0$ be fixed and denote $U = A(\sum_{k=1}^n \varepsilon_k / \sqrt{n})$ and $V = (Q_n + R_n) / \sqrt{n}$. So $U + V = \sum_{k=1}^n X_k / \sqrt{n}$ and

$$\mathbb{P}(\|U + V\|_H \leq t) \leq \mathbb{P}(\|U\|_H \leq t + \lambda) + \mathbb{P}(\|V\|_H \geq \lambda) \quad (5)$$

For $\lambda_0 = 2\|V\|_\infty$, we obtain

$$\mathbb{P}(\|U + V\|_H \leq t) - \mathbb{P}(\|N\|_H \leq t) \leq \mathbb{P}(\|U\|_H \leq t + \lambda_0) - \mathbb{P}(\|N\|_H \leq t).$$

If $c(N)$ denotes a bound for the density of $\|N\|_H$ (see Davydov et al. [5]) then

$$\Delta_n(X) \leq \Delta_n(A(\varepsilon)) + \frac{2c(N)\|Q_n + R_n\|_\infty}{\sqrt{n}}.$$

Noting that

$$Q_n = \sum_{j \geq n+2} a_j \left(\sum_{k=1-j}^{-n-1} \varepsilon_k \right) + \sum_{j < 0} a_j \left(\sum_{k=n+1}^{n-j} \varepsilon_k \right) \quad (6)$$

and

$$R_n = R'_n + R''_n \quad (7)$$

where

$$R'_n = - \sum_{j=-n}^{-1} a_j \left(\sum_{k=1}^{-j} \varepsilon_k \right) - \sum_{j < -n} a_j \left(\sum_{k=1}^n \varepsilon_k \right) - \sum_{j > 0} a_j \left(\sum_{k=n-j+1}^n \varepsilon_k \right)$$

and

$$R_n'' = \sum_{j=1}^n a_j \left(\sum_{k=-j+1}^0 \varepsilon_k \right) + \sum_{j=n+1}^{2n} a_j \left(\sum_{k=-n}^{n-j} \varepsilon_k \right),$$

we derive that $\|Q_n + R_n\|_\infty \leq 7\|\varepsilon_0\|_\infty \sum_{j \in \mathbb{Z}} |j| \|a_j\|_{\mathcal{L}(H)}$ and consequently

$$\Delta_n(X) \leq \Delta_n(A(\varepsilon)) + \frac{14c(N)\|\varepsilon_0\|_\infty \sum_{j \in \mathbb{Z}} |j| \|a_j\|_{\mathcal{L}(H)}}{\sqrt{n}}.$$

Combining the last inequality with the Berry-Esseen inequality for i.i.d. centered H -valued random variables (see Yurinski [11] or Bosq [2], Theorem 2.9) we obtain (3).

In the other part, if ψ is a Young function we have $\mathbb{P}(\|V\|_H \geq \lambda) \leq \frac{1}{\psi(\lambda/\|V\|_\psi)}$ and keeping in mind inequality (5), we derive

$$\Delta_n(X) \leq \Delta_n(A(\varepsilon)) + c(N)\lambda + \frac{1}{\psi(\lambda/\|V\|_\psi)}.$$

Noting that $c(N)\lambda = \frac{1}{\psi(\lambda/\|V\|_\psi)}$ if and only if $\lambda = \frac{\varphi(c(N)\|V\|_\psi)}{c(N)}$ where φ is defined by $\varphi(x) = xh^{-1}(1/x)$ and h by $h(x) = x\psi(x)$, we conclude

$$\Delta_n(X) \leq \Delta_n(A(\varepsilon)) + \varphi\left(\frac{c(N)\|Q_n + R_n\|_\psi}{\sqrt{n}}\right).$$

The proof of Theorem 1 is complete.

Proof of Corollary 1. Assume that $\|\varepsilon_0\|_{\psi_1} < \infty$ where ψ_1 is the Young function defined by $\psi_1(x) = \exp(x) - 1$. There exists $a > 0$ such that $E(\exp(a\|\varepsilon_0\|_H)) \leq 2$. So, there exist (see Arak and Zaizsev [1]) constants B and L such that

$$E\|\varepsilon_0\|_H^m \leq \frac{m!}{2} B^2 L^{m-2}, \quad m = 2, 3, 4, \dots$$

Applying Pinelis-Sakhanenko inequality (see Pinelis and Sakhanenko [?] or Bosq [2]), we obtain

$$\mathbb{P}\left(\left\|\sum_{k=p}^q \varepsilon_k\right\|_H \geq x\right) \leq \exp\left(-\frac{x^2}{2(q-p+1)B^2 + 2xL}\right), \quad x > 0$$

and using Lemma 2.2.10 in Van Der Vaart and Wellner [?], there exists a universal constant K such that

$$\left\|\sum_{k=p}^q \varepsilon_k\right\|_{\psi_1} \leq K\left(L + B\sqrt{q-p+1}\right) \quad (8)$$

Combining (6), (7) and (8), we derive $\|Q_n + R_n\|_{\psi_1} \leq C \sum_{j \in \mathbb{Z}} \sqrt{|j|} \|a_j\|_{\mathcal{L}(H)}$ where the constant C does not depend on n . Keeping in mind the Berry-Esseen's central limit theorem for i.i.d. centered H -valued random variables (see Yurinski [11] or Bosq [2], Theorem 2.9), we apply Theorem 1 with the Young function ψ_1 . Since the function φ defined by $\varphi(x) = xh^{-1}(1/x)$ with $h(x) = x\psi_1(x)$ satisfies

$$\lim_{x \rightarrow 0} \frac{\varphi(x)}{x \log(1 + \frac{1}{x})} = 0,$$

we derive $\Delta_n(X) = O\left(\frac{\log n}{\sqrt{n}}\right)$.

Now, assume that $\|\varepsilon_0\|_r < \infty$ for some $r \geq 3$. Applying Pinelis inequality (see Pinelis [8]), there exists a universal constant K such that

$$\left\| \sum_{k=p}^q \varepsilon_k \right\|_r \leq K \left(r \left(\sum_{k=p}^q E \|\varepsilon_k\|_H^r \right)^{1/r} + \sqrt{r} \left(\sum_{k=p}^q E \|\varepsilon_k\|_H^2 \right)^{1/2} \right)$$

and consequently

$$\left\| \sum_{k=p}^q \varepsilon_k \right\|_r \leq 2Kr \|\varepsilon_0\|_r \sqrt{q - p + 1}. \quad (9)$$

Combining (6), (7) and (9), we derive $\|Q_n + R_n\|_r \leq C \sum_{j \in \mathbb{Z}} \sqrt{|j|} \|a_j\|_{\mathcal{L}(H)}$ where the constant C does not depend on n . Again, applying Berry-Esseen's CLT (see Yurinski [11] or Bosq [2], Theorem 2.9) and Theorem 1 with the Young function $\psi(x) = x^r$ and the function φ given by $\varphi(x) = x^{r/(r+1)}$, we obtain $\Delta_n(X) = O\left(n^{-\frac{r}{2(r+1)}}\right)$. The proof of Corollary 1 is complete.

References

- [1] Arak, T. V. and Zaitsev, A. Yu., *Uniform limit theorems for sums of independent random variables*, Proc. Steklov Inst. Math. **174** (1988).
- [2] Bosq, D., *Linear processes in function spaces*, Lecture Notes in Statistics, **149**, Springer-Verlag, 2000.
- [3] Bosq, D., Berry-Esseen inequality for linear processes in Hilbert spaces, *Statist. Probab. Lett.*, **63** (3), (2003), 243–247.

- [4] Bosq, D., Erratum and complements to: “Berry-Esseen inequality for linear processes in Hilbert spaces” [Statist. Probab. Lett. **63** (2003), no. 3, 243–247], *Statist. Probab. Lett.*, **70** (2), (2004), 171–174.
- [5] Davydov, Yu. A. and Lifshits, M. A. and Smorodina, N. V., *Local properties of distributions of stochastic functionals*, **173**, American Mathematical Society, Providence, RI 1998.
- [6] Merlevède, F., Peligrad, M. and Utev, S., Sharp conditions for the CLT of linear processes in a Hilbert space, *J. Theoret. Probab.*, **10** (3), (1997), 681–693.
- [7] Paulauskas, V., Personal communication, 2004.
- [8] Pinelis, I., Optimum bounds for the distributions of martingales in Banach spaces, *Ann. Probab.*, **22** (4), (1994), 1679–1706.
- [9] Pinelis, I. F. and Sakhanenko, A. I., Remarks on inequalities for probabilities of large deviations, *Teor. Veroyatnost. i Primenen.*, **30** (1), (1985), 127–131.
- [10] van der Vaart, Aad W. and Wellner, Jon A., *Weak convergence and empirical processes*, Springer Series in Statistics, Springer-Verlag, New York, 1996.
- [11] Yurinskii, V. V., On the accuracy of normal approximation of the probability of hitting a ball, *Teor. Veroyatnost. i Primenen.*, **27** (2), (1982), 270–278.

Mohamed EL MACHKOURI
 Laboratoire de Mathématiques Raphaël Salem
 UMR CNRS 6085, Université de Rouen
 Avenue de l’université
 76801 Saint-Etienne du Rouvray
 mohamed.elmachkouri@univ-rouen.fr